

Query Match	9.6%	Score 146.6	DB 4	Length 876
Best Local Similarity	52.28	Pred. No. 1.7e-41		
Matches 385	Conservative	0	Mis-matches 334	Indels 18
				Gaps 2
QY	ATTATTAAGCTGGGAGCTGGAGCTGTATGACAGTCAACACGACCAAGTTGTGTGCCTGAC	870		
DB	136 ATTTTACATGGGATTTCACACTTTATGACCTTTCAAAGGTTACTTAAAGTGGGTTTAAAC	195		
QY	871 AAGACCTGATTTTGTGCTGTGCATGACATTAAGCTCTGCTGTATCCGCTCAGGAA	930		
DB	196 AAAGATTTGTTTGTTCCTCCAAAGTTGACGATGAGTTGTCAATTTGCTGCTTTAAT	255		
QY	931 GCTCTGCTGGCTCATCCG-----ACAGTACTTCACTAGCTCTATCAAGATGTCGT	984		
DB	256 GCATTGATGATTCTACCGTGGACAACTATTTGGCCGAGATTTCTTTTCAATTTGTGGG	315		
QY	985 ATGTTTGAATGACGAGAAATTGGAAGCTGTGGCCGAGAGCTGCATCCAACTTCATG	1044		
DB	316 TTGTTTGAATGAGAAATTGGCTCTCTCACTACAGAAAGTGTCTAGAGGTGAATTAAT	375		
QY	1045 AGCAGTGCATAGAGCTATTACGGAACTTTCACCCAAATTACGCTCTTACGTCGTG	1104		
DB	376 GAATTAGTTGTGACAGAGTGTTCATCTTAATTTCTTAATTCAGGAAGTTCTTGATATT	435		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 03:55:19 ; Search time 961 Seconds  
(without alignments)  
10242.485 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602  
Perfect score: 1530  
Sequence: 1 atgacccaagaagatgtctt.....acaagaattgcagatttc 1530

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/prodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/prodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/prodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/prodata/1/pubpna/US10I\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/prodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/prodata/1/pubpna/US11A\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	1679	20	US-10-664-958-4 Sequence 4, Appl1
2	787.6	51.5	1916	20	US-10-664-958-2 Sequence 2, Appl1
3	787.6	51.5	3383	20	US-10-664-958-1 Sequence 1, Appl1
4	184.6	12.1	707	20	US-10-653-047-448 Sequence 448, App
5	119.2	7.8	1656	18	US-10-425-114-11096 Sequence 11096, A
6	117.6	7.7	1807	18	US-10-425-114-3262 Sequence 3262, Ap
7	117.6	7.7	1948	20	US-10-425-115-5515 Sequence 5515, Ap

8	115.2	7.5	844	20	US-10-425-115-42258 Sequence 42258, A
9	109.2	7.1	1696	17	US-10-353-690-69 Sequence 69, Appl1
10	109.2	7.1	1696	21	US-10-956-157-1776 Sequence 1776, Ap
11	109.2	7.1	2583	20	US-10-357-930-22225 Sequence 22225, A
12	109.2	7.1	2583	20	US-10-357-930-22225 Sequence 22225, A
13	108	7.1	1975	18	US-10-424-599-14244 Sequence 14244, A
14	95.4	6.2	1784	19	US-10-437-963-98779 Sequence 98779, A
15	86.6	5.7	3380	15	US-10-128-714-416 Sequence 416, App
16	86.6	5.7	3381	15	US-10-128-714-5416 Sequence 5416, Ap
17	83.8	5.5	1117	19	US-10-767-792-74 Sequence 74, Appl1
18	82	5.4	844	20	US-10-425-115-54577 Sequence 54577, A
19	79.6	5.2	1284	18	US-10-148-414-8 Sequence 8, Appl1
20	78.2	5.1	637	19	US-10-767-701-6053 Sequence 6053, Ap
21	76.8	5.0	600	21	US-10-956-157-7011 Sequence 7011, Ap
22	76.4	5.0	681	19	US-10-283-978A-763 Sequence 763, App
23	70.2	4.6	544	20	US-10-425-115-82175 Sequence 82175, A
24	70	4.6	754	19	US-10-767-701-6113 Sequence 6113, Ap
25	69.2	4.5	1576	18	US-10-425-114-29546 Sequence 29546, A
26	66.6	4.4	607	19	US-10-021-323-4253 Sequence 4253, Ap
27	63.4	4.1	215	18	US-10-424-599-140276 Sequence 140276, A
28	62.6	4.1	489	11	US-09-732-627A-3018 Sequence 3018, Ap
29	61.4	4.0	583	20	US-10-425-115-101845 Sequence 101845, A
30	59.6	3.9	414	10	US-09-918-995-3652 Sequence 3652, Ap
31	59.6	3.9	421	10	US-09-918-995-4180 Sequence 4180, Ap
32	58.8	3.8	735	19	US-10-767-701-26797 Sequence 26797, A
33	54.8	3.6	365	10	US-09-918-995-8281 Sequence 8281, Ap
34	51.8	3.4	442	20	US-10-357-930-32748 Sequence 32748, A
35	50	3.3	571	19	US-10-021-323-42609 Sequence 42609, Ap
36	49.2	3.2	827	18	US-10-424-599-140279 Sequence 140279, A
37	48.8	3.2	463	20	US-10-357-930-33748 Sequence 33748, A
38	48.6	3.2	463	20	US-10-357-930-33748 Sequence 33748, A
39	48.6	3.2	463	20	US-10-357-930-33748 Sequence 33748, A
40	45.4	3.0	1236	15	US-10-156-761-4373 Sequence 4373, Ap
41	45.4	3.0	9025608	17	US-10-156-761-1 Sequence 1, Appl1
42	43.8	2.9	3662	15	US-10-320-797-205 Sequence 205, Appl1
43	43.4	2.8	607	14	US-10-123-155-102 Sequence 102, App
44	43.4	2.8	607	15	US-10-146-731-102 Sequence 102, App
45	43.4	2.8	607	15	US-10-140-472-102 Sequence 102, App

#### ALIGNMENTS

RESULT 1  
US-10-664-958-4  
Sequence 4, Application US/10664958  
Publication No. US20040219636A1  
GENERAL INFORMATION:  
APPLICANT: KOIBUCHI, KYOKO  
APPLICANT: NINOMIYA, DAIKI  
APPLICANT: KOTIMA, MARI  
APPLICANT: UEDA, YOICHI  
APPLICANT: MATSUMOTO, JUN-ICHI  
APPLICANT: KITAMOTO, KATSUHIKO  
TITLE OF INVENTION: NEW AMINOPEPTIDASE AND THE GENES THEREOF  
FILE REFERENCE: 241461USOCONT  
CURRENT APPLICATION NUMBER: US/10/664, 958  
CURRENT FILING DATE: 2003-09-22  
PRIOR APPLICATION NUMBER: PCT/JP02/02476  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: JP 2001-078930  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: JP 2001-293348  
PRIOR FILING DATE: 2001-09-26  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 4  
LENGTH: 1679  
TYPE: DNA  
ORGANISM: Aspergillus oryzae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (73)..(1602)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 01:48:04 ; Search time 4629 Seconds

(without alignments)  
12581.191 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602

Sequence: 1 atgaccagaagagtgctcct.....acaaggaatttcgacatttc 1530

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST.\*  
1: gb\_ests1.\*  
2: gb\_ests2.\*  
3: gb\_hnc.\*  
4: gb\_ests3.\*  
5: gb\_ests4.\*  
6: gb\_ests5.\*  
7: gb\_ests6.\*  
8: gb\_g881.\*  
9: gb\_g882.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482	31.5	636	7	CO150272 EST825325
2	453.4	29.6	684	7	CO136018 EST830689
3	419.8	27.4	898	7	CO005654 EST793989
4	409.4	26.8	901	7	CO013013 EST801348
5	382.6	25.0	827	7	CO005655 EST793990
6	352	23.0	780	7	CO009866 EST798201
7	320.8	21.0	859	7	CO013014 EST801349
8	316.4	20.7	535	1	AT211309 O6D09a1.r
9	277.4	18.1	530	1	AT211331 O6D03a1.r
10	256.4	16.8	764	6	CD485799 CF083.1H0
11	247.6	16.2	759	6	CB903053 trlC032x1
12	247.4	16.2	691	7	CF873032 trlC032x1
13	231	15.1	768	9	CL558944 OB_Ba001
14	215	14.1	391	5	BQ499748 EST08973
15	211.2	13.8	828	2	BE643133
16	210.2	13.7	434	7	CN249799 EST015718
17	195.6	12.8	539	1	AT211330 O6D03a1.f
18	184.6	12.1	1030	7	CN248531 EST014438
19	171.6	11.2	758	7	CN240819 EST006682
20	167.4	10.9	915	7	CF883173 trlC032x1
21	159.2	10.4	519	5	BQ503674 EST12904
22	154.4	10.1	734	5	CF883059 trlC032x1
23	153	10.0	571	9	CL559297 OB_Ba001
24	153	10.0	580	7	CN249701 EST015620

25	143.4	9.4	463	2	BF252877	BF252877 EST420140
26	140.2	9.2	974	7	CNS06Y0Z	AL421305 T3 end of
27	136.8	8.9	268	7	CO009867	CO009867 EST798202
28	133.8	8.7	441	5	BQ502774	BQ502774 EST12004
29	126.2	8.2	641	7	CK907937	CK907937 rhzma0.00
30	121	7.9	780	7	CO019776	CO019776 EST815869
31	118.2	7.7	584	8	BZ293586	BZ293586 CG0412.r1
32	114.4	7.5	1422	9	AY399515	AY399515 Mus muscu
33	112.8	7.4	2489	9	AK038646	AK038646 Mus muscu
34	111.8	7.3	885	7	CF821438	CF821438 EST69820
35	109.2	7.1	1509	9	AY399513	AY399513 Homo sapi
36	109.2	7.1	1560	3	CNS0A1EQ	BX829510 Arabidops
37	107.8	7.0	2835	3	AK052001	AK052001 Mus muscu
38	107.6	7.0	1319	3	CR602255	CR602255 full-1eng
39	107.6	7.0	1527	3	CR601156	CR601156 full-1eng
40	107.6	7.0	1528	3	CR620845	CR620845 full-1eng
41	107.6	7.0	1544	3	CR605917	CR605917 full-1eng
42	107.6	7.0	1593	3	CR608875	CR608875 full-1eng
43	107.6	7.0	1594	3	CR593910	CR593910 full-1eng
44	107.6	7.0	1602	3	CR607064	CR607064 full-1eng
45	107.6	7.0	1622	3	CR611309	CR611309 full-1eng

#### ALIGNMENTS

RESULT 1  
LOCUS CO150272 636 bp mRNA linear EST 17-JUN-2004  
DEFINITION EST825325 Aspergillus flavus Normalized cDNA Expression Library  
Aspergillus flavus cDNA clone NAGEB35 5' end, mRNA sequence.  
ACCESSION CO150272  
VERSION CO150272.1 GI:48904273  
KEYWORDS EST.

#### SOURCE

ORGANISM Aspergillus flavus

Aspergillus flavus  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

#### REFERENCE

AUTHORS Yu, J., White, C.A., Nieman, W.C., Bhattacharya, D. and Cleveland, T.B.

TITLE Aspergillus flavus expressed sequence tags for identification of

genes with putative roles in aflatoxin contamination of crops

FEMS Microbiol. Lett. (2004) in press

#### JOURNAL

COMMENT

Contact: Yu J

Food and Feed Safety Research Unit

USDA/ARS, Southern Regional Research Center

1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA

Tel: 504 286 4405

Fax: 504 286 4419

Email: jlyu@arcc.ars.usda.gov

Contact Dr. Yu at USDA/ARS SRRC (jlyu@arcc.ars.usda.gov) for clone

information

PCR Primers

FORWARD: M13F

BACKWARD: M13R

Seq primer: M13 Forward

POLYA-No.

#### FEATURES

SOURCE

Location/Qualifiers

1..636

/organism="Aspergillus flavus"

/mol\_type="mRNA"

/strain="NRRL 3357"

/db\_xref="taxon:5059"

/clone="NAGEB35"

/sex="asexual mycelia"

/cell\_type="mycelia"

/dev\_stage="developmental stages from 18 to 96 hours"

/lab\_host="E. coli DH10B T1 resistant cells"

/clone\_lib="Aspergillus flavus Normalized cDNA Expression

Library"

/note="Vector: pBlueScript (SK+) (Stratagene), antibiotic

selection marker: Carbenicillin; Site 1: NotI, at the 5

prime end; Site 2: BcoRI, at the 3 prime end; This

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 01:45:05 ; Search time 6488 Seconds  
(without alignments)  
11426.708 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602  
Perfect score: 1530  
Sequence: 1 atgacccaagaagtgctcct.....acaagaatttcgacatttc 1530

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_srs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	1679	6	BD178569
2	1530	100.0	1679	6	BD178569
3	1530	100.0	1679	6	BD178569
4	1530	100.0	1679	6	BD178569
5	1530	100.0	1679	6	BD178569
6	1530	100.0	1679	6	BD178569
7	1530	100.0	1679	6	BD178569
8	1530	100.0	1679	6	BD178569
9	1530	100.0	1679	6	BD178569
10	1530	100.0	1679	6	BD178569
11	1530	100.0	1679	6	BD178569
12	1530	100.0	1679	6	BD178569
13	1530	100.0	1679	6	BD178569
14	1530	100.0	1679	6	BD178569
15	1530	100.0	1679	6	BD178569
16	1530	100.0	1679	6	BD178569
17	1530	100.0	1679	6	BD178569
18	1530	100.0	1679	6	BD178569
19	1530	100.0	1679	6	BD178569

20	149	9.7	110000	8	CR382125_18	Continuation (19 o
21	149	9.7	110000	8	CR382125_19	Continuation (20 o
22	146.6	9.6	876	8	ARS50239	ARS50239 Sequence
23	146.6	8.9	780	8	CMS01CDM	ALA15394 Botrytis
24	127.8	8.4	1758	8	AF420303	AF420303 Coccidiol
25	123.2	8.1	110000	8	CR380955_00	CR380955 Candida g
26	123.2	8.1	110000	8	CR380955_01	Continuation (2 of
27	120.2	7.9	1515	5	AJ720587	AJ720587 Gallus ga
28	120.2	7.9	1595	9	BC003040	BC003040 Homo sapi
29	119.4	7.8	39954	8	YSC8263	U00059 Saccharomyc
30	114.4	7.5	1637	6	AR127148	AR127148 Sequence
31	114.4	7.5	1637	6	E36332	E36332 METPRO 02 b
32	114.4	7.5	1637	6	AX003001	AX003001 Sequence
33	112.8	7.4	1777	10	AF005051	AF005051 Mus muscu
34	111	7.3	110000	8	CR382138_16	Continuation (17 o
35	109.2	7.1	1637	6	AR127146	AR127146 Sequence
36	109.2	7.1	1637	6	E36330	E36330 METPRO 02 b
37	109.2	7.1	1637	6	AX002997	AX002997 Sequence
38	109.2	7.1	1671	6	BD156930	BD156930 Primer fo
39	109.2	7.1	1671	6	AX877932	AX877932 Sequence
40	109.2	7.1	1671	9	AK001977	AK001977 Homo sapi
41	109.2	7.1	1696	9	AF005050	AF005050 Homo sapi
42	109.2	7.1	1754	9	BC000653	BC000653 Homo sapi
43	109.2	7.1	1910	6	AR123707	AR123707 Sequence
44	109.2	7.1	1910	6	BD205508	BD205508 Human pro
45	109.2	7.1	2583	6	CQ490358	CQ490358 Sequence

## ALIGNMENTS

RESULT 1	BD178569	1679 bp	DNA	linear	PAT 16-APR-2003
LOCUS	BD178569				
DEFINITION	Novel aminopeptidase and its gene.				
ACCESSION	BD178569.1	GI:30015835			
VERSION	WO 02077223-A/3.				
KEYWORDS	Aspergillus oryzae				
SOURCE	Aspergillus oryzae				
ORGANISM	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
REFERENCE	1 (bases 1 to 1679)				
AUTHORS	Kojima, K., Ninomiya, D., Kojima, M., Ueda, Y., Maruyama, J. and Kitamoto, K.				
TITLE	Novel aminopeptidase and its gene				
JOURNAL	Patent: WO 02077223-A 3 03-OCT-2002; AJINOMOTO CO INC, KYOKO KOIBUCHI, DAIKI NINOMIYA, MARI KOJIMA, YOICHI UEDA, JUNICHI MARUYAMA, KATSUHIKO KITAMOTO				
COMMENT	OS Aspergillus oryzae PN WO 02077223-A/3 PD 03-OCT-2002 PF 15-MAR-2002 WO 2002JP02476 PR 19-MAR-2001 JP 01P 078930, 26-SEP-2001 JP 01P 293348 PI KYOKO KOIBUCHI, DAIKI NINOMIYA, MARI KOJIMA, YOICHI UEDA, JUNICHI MARUYAMA, KATSUHIKO KITAMOTO				
FEATURES	source	1..1679			
	Location/Qualifiers	/organism="Aspergillus oryzae"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:5062"			
ORIGIN	Query Match	100.0%; Score 1530; DB 6; Length 1679;			
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches 1530; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGACCAAAAGAGTGCTTATCTCGGTATTCGCGATGCGATGCGCTTATGCGCGTGGCC	60		

xx  
ps Claim 1; Page 45-49; 61pp; Japanese..

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 11, 2005, 15:11:49; Search time 125.5 seconds  
(without alignments)  
9617.214 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602

Perfect score: 2784  
Sequence: 1 atgaccaaagagtgctct.....acaggaatttcagatttc 1530

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1777461 seqs, 394431504 residues

Total number of hits satisfying chosen parameters: 3554922

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US10664958/funat\_10092005.173357\_2052/app.query.fasta.1.1671  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=top20.rapb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human0.cdi -LIST=45 -BOCLALIGN=200 -THR SCORE=Pct -THR MAX=100  
-THR MIN=0 -ALIGN=20 -MODE=LOCAL -OUTPMT=pct -NORM=ext -HAPBITZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10664958.@CGN\_1\_130@funat\_10092005.173357\_2052  
-NCPU=6 -ICPU=3 -NO MMAP -LARGESOURCY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database :  
Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppa/US10E\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppa/US10F\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2626	94.3	510	16	US-10-664-958-5
2	2112.5	75.9	519	16	US-10-664-958-3
3	846.5	30.4	496	15	US-10-425-114-44595
4	831.5	29.5	565	15	US-10-425-114-47233
5	821.5	29.5	487	15	US-10-425-114-47233
6	814	29.2	475	16	US-10-425-115-190178
7	806.5	29.0	478	16	US-10-437-963-201262
8	737.5	26.5	425	15	US-10-353-690-70
9	676	24.3	478	15	US-10-148-414-9
10	553.5	19.9	491	15	US-10-425-599-283118
11	498.5	17.9	420	9	US-09-804-073-2
12	498.5	17.9	441	9	US-09-738-626-5156
13	486.5	17.5	432	14	US-10-156-761-11923
14	434.5	15.6	253	16	US-10-425-115-239240
15	406	14.6	240	16	US-10-425-115-226921
16	352	12.6	161	16	US-10-767-701-37617
17	349.5	12.6	258	16	US-10-425-114-55505
18	310.5	11.2	226	16	US-10-767-701-58361
19	300.5	10.8	173	16	US-10-767-701-37677
20	276.5	9.9	155	16	US-10-425-115-286508
21	268.5	9.6	181	16	US-10-425-115-268838
22	266.5	9.6	160	16	US-10-437-963-181727
23	254	9.1	179	16	US-10-767-701-36045
24	250	9.0	138	15	US-10-424-599-283121
25	244.5	8.8	154	16	US-10-437-963-116670
26	228	8.2	163	16	US-10-437-963-182169
27	192.5	6.9	165	16	US-10-437-963-120717
28	179.5	6.4	114	16	US-10-437-963-120717
29	148.5	5.3	97	15	US-10-424-599-274525
30	135	4.8	65	16	US-10-425-115-308560
31	121	4.3	905	15	US-10-108-260A-3889
32	120.5	4.3	925	15	US-10-108-260A-3889
33	120	4.3	528	10	US-09-840-746-20
34	118	4.2	519	15	US-10-369-493-22438
35	118	4.2	519	16	US-10-451-467A-774
36	117	4.2	19723	15	US-10-084-846A-5
37	116	4.2	19695	15	US-10-084-846A-3
38	115.5	4.1	463	16	US-10-437-963-109124
39	115	4.1	487	16	US-10-437-963-117896
40	111.5	4.0	6842	15	US-10-461-194-131
41	111	4.0	2442	13	US-10-109-886-10
42	111	4.0	2442	16	US-10-473-127-631
43	111	4.0	2442	16	US-10-473-127-631
44	111	4.0	2442	16	US-10-473-127-645
45	111	4.0	2442	16	US-10-473-127-647

ALIGNMENTS

RESULT 1  
US-10-664-958-5  
Sequence 5, Application US/10664958  
Publication No. US20040219636A1  
GENERAL INFORMATION:  
APPLICANT: KOJIBUCHI, KYOKO  
APPLICANT: NINOMIYA, DAIKI  
APPLICANT: KOJIMA, MARI  
APPLICANT: UEDA, YOICHI  
APPLICANT: MARUYAMA, JUN-ICHI  
APPLICANT: KITAMOTO, KATSUHIKO  
TITLE OF INVENTION: NEW AMINOPEPTIDASE AND THE GENES THEREOF  
FILE REFERENCE: 241461US0CONT  
CURRENT APPLICATION NUMBER: US/10/664,958  
CURRENT FILING DATE: 2003-09-22  
PRIOR APPLICATION NUMBER: PCT/JP02/02476  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: JP 2001-078930  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: JP 2001-293348

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 11, 2005, 14:45:13 ; Search time 37 Seconds  
(without alignments)  
6173.679 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602  
Perfect score: 2784  
Sequence: 1 atgaccacaaagagtgctc.....acaaggaattgcagatttc 1530

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp  
-O=/cgn2\_1/USPFO\_appool\_p/US10664958/r/nat.10092005.173356.1993/app\_query.fasta\_1.1671  
-DB=Issued\_Patents\_AA -QFRT=fastan -SUFFIX=top20.ral -MINMATCH=0.1 -LOOPTL=0  
-LOOEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10664958 @cgn 1.1.33 @nat.10092005.173356.1993 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents AA:

1: /cgn2\_6/ptodata/1/1aa/5A COMB.pdp:\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pdp:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pdp:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pdp:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS COMB.pdp:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1013.5	36.4	514	4	US-09-538-092-495 Sequence 495, App
2	811.5	28.1	487	3	US-09-525-046-2 Sequence 2, Appl
3	737.5	26.5	471	3	US-09-071-709-4 Sequence 4, Appl
4	737.5	26.5	471	3	US-09-160-494-2 Sequence 2, Appl
5	732.5	26.3	471	3	US-09-160-494-6 Sequence 6, Appl
6	681.5	24.5	485	4	US-09-252-991A-21636 Sequence 21636, A
7	645.5	23.2	291	4	US-09-248-796A-19473 Sequence 19473, A
8	640	23.0	470	3	US-09-071-709-12 Sequence 12, Appl
9	559.5	20.1	459	3	US-09-160-494-4 Sequence 4, Appl
10	542.5	19.5	447	4	US-08-311-731A-246 Sequence 246, App
11	498.5	17.9	420	4	US-09-602-777A-110 Sequence 110, App
12	473	17.0	291	4	US-09-248-796A-19474 Sequence 19474, A

13	451.5	16.2	318	4	US-09-248-796A-19475 Sequence 19475, A
14	448.5	16.1	249	4	US-09-248-796A-19487 Sequence 19487, A
15	412	14.8	225	4	US-09-248-796A-19486 Sequence 19486, A
16	244.5	8.8	283	4	US-09-248-796A-19476 Sequence 19476, A
17	141	5.1	752	4	US-09-252-991A-21724 Sequence 21724, A
18	119	4.3	1093	4	US-09-252-991A-21827 Sequence 21827, A
19	112.5	4.0	629	4	US-09-252-991A-17150 Sequence 17150, A
20	111	4.0	2442	3	US-09-514-247A-10 Sequence 1370, A
21	111	4.0	2442	4	US-09-538-092-1370 Sequence 16, Appl
22	109	3.9	648	4	US-10-263-687A-16 Sequence 16, Appl
23	109	3.9	648	4	US-10-263-687A-16 Sequence 16, Appl
24	109	3.9	1220	1	US-08-158-232-43 Sequence 43, Appl
25	109	3.9	1220	2	US-08-611-928-43 Sequence 43, Appl
26	109	3.9	1220	3	US-09-173-891-43 Sequence 43, Appl
27	108.5	3.9	605	2	US-08-889-402-2 Sequence 2, Appl
28	107.5	3.9	2090	4	US-09-538-092-1081 Sequence 1081, Ap
29	107.5	3.9	2120	4	US-09-949-016-9768 Sequence 9768, Ap
30	106.5	3.8	654	4	US-09-437-687A-2 Sequence 2, Appl
31	106.5	3.8	654	4	US-10-263-250-2 Sequence 2, Appl
32	105.5	3.8	620	2	US-08-419-652-7 Sequence 7, Appl
33	105.5	3.8	898	1	US-08-465-995A-4 Sequence 4, Appl
34	105.5	3.8	898	2	US-08-465-994C-4 Sequence 4, Appl
35	105.5	3.8	898	2	US-08-966-145-4 Sequence 4, Appl
36	105.5	3.8	920	1	US-08-101-593-4 Sequence 6, Appl
37	105.5	3.8	1001	1	US-07-797-556-6 Sequence 6, Appl
38	105.5	3.8	1001	1	US-07-943-843-2 Sequence 2, Appl
39	105.5	3.8	1001	1	US-08-347-003-2 Sequence 2, Appl
40	105.5	3.8	1097	1	US-07-943-843-6 Sequence 6, Appl
41	105.5	3.8	1097	1	US-08-347-003-6 Sequence 6, Appl
42	105.5	3.8	1097	1	US-08-949-016-6209 Sequence 6209, Ap
43	105.5	3.8	1119	4	US-09-949-016-10081 Sequence 10081, A
44	105.5	3.8	1385	4	US-09-252-991A-21919 Sequence 21919, A
45	104.5	3.8	898	1	US-08-465-995A-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-538-092-495  
Sequence 495, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Iolc  
APPLICANT: Mansfield, Treci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
PRIORITY FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuratSeqFormatter Version 0.9  
SEQ ID NO 495  
LENGTH: 514  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number YK1403C  
US-09-538-092-495

Alignment Scores:  
Pred. No.: 8.21e-96  
Score: 1013.50  
Percent Similarity: 60.04%  
Best Local Similarity: 42.44%  
Query Match: 36.40%  
DB: 4  
Gaps: 8

US-10-664-958-4\_COPY\_73\_1602 (1-1530) \* US-09-538-092-495 (1-514)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 11, 2005, 05:18:03 ; Search time 158.5 Seconds  
(without alignments)  
9886.198 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602

Perfect score: 2784  
Sequence: 1 atgacccaagaagagtgctc.....acaaggaattgcagatttc 1530

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp  
-Q=/cgn2\_1/USFO.spool\_p/US1066958/runat.10092005.173355.1973/app\_query.fasta.1.1671  
-DB=uniprot\_03 -QFMT=fastan -SUPFIX=top20.rup -MINMATCH=0.1 -LOOPT=0  
-LIST=45 -OCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US1066958\_QCN 1.1 153 @runat.10092005.173355.1973 -NCP=6 -ICP=3  
-NO\_MAP -LARGE=0 -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Uniprot 03:\*

1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1528	54.9	535	2	Q9H85	Q9H85 neurospora
2	1174	42.2	493	2	O6CCD1	O6CCD1 yarrowia 11
3	1040	37.4	509	2	O6FMF2	O6FMF2 candida gla
4	1034.5	37.2	519	2	O6CK48	O6CK48 kluyveromyc
5	1013.5	36.4	514	1	AMPL_YEAST	PL4904 saccharomyc
6	1005	36.1	506	2	O6BHT6	O6BHT6 debaryomyc
7	998	35.8	469	2	O6C022	O6C022 yarrowia 11
8	907.5	32.6	498	2	O75DQ2	O75DQ2 ashyia goss
9	883.5	31.7	476	2	O6CFM9	O6CFM9 yarrowia 11
10	867	31.1	495	2	O6CTK5	O6CTK5 coccidioid
11	866.5	31.1	499	2	O6CMCS	O6CMCS kluyveromyc
12	862.5	31.0	523	2	O6BKU6	O6BKU6 debaryomyc
13	857	30.8	493	2	O6BHZ6	O6BHZ6 debaryomyc
14	856	30.7	484	2	O6FRJ3	O6FRJ3 candida gla
15	843	30.3	490	1	DNPE_YEAST	O7RYK7 neurospora
16						P38821 saccharomyc

17	814.5	29.3	467	1	DNPE_SCHPO	O36014 schizosacch
18	782.5	28.1	477	2	O9LSTO	O9LSTO arabidopsis
19	777.5	27.9	477	2	O8LFF32	O8LFF32 arabidopsis
20	737.5	26.5	475	1	DNPE_HUMAN	O9U140 homo sapien
21	737	26.5	525	2	O8SA06	O8SA06 oryza sativ
22	728.5	26.2	473	1	DNPE_MOUSE	O922W0 mus musculu
23	728.5	26.2	473	2	O8CAJ7	O8CAJ7 mus musculu
24	718	25.8	526	2	O9LZ26	O9LZ26 arabidopsis
25	714	25.6	526	2	O8GWI9	O8GWI9 arabidopsis
26	700.5	25.2	470	2	O8BPM9	O8BPM9 mus musculu
27	699.5	25.1	469	2	O8Q3B5	O8Q3B5 xenopus lae
28	692	24.9	479	2	O6Q061	O6Q061 xenopus lae
29	682	24.5	429	1	APER_PSEPK	O88M44 pseudomonas
30	677	24.3	429	1	APER_PSEAR	O9HY23 pseudomonas
31	670	24.1	429	1	APER_PSESM	O9HY23 pseudomonas
32	640	23.0	433	1	APER_CLOAB	O9L7F4 clostridium
33	640	23.0	470	1	DNPE_CABSL	O13087 caenorhabdi
34	627.5	22.5	430	2	O89IC3	O89IC3 clostridium
35	613.5	22.0	441	2	O9NV55	O9NV55 homo sapien
36	605	21.7	437	2	O6AL63	O6AL63 desulfotale
37	582	20.9	430	2	O73PHO	O73PHO treponema d
38	581	20.9	431	2	O8XMT1	O8XMT1 clostridium
39	560.5	20.1	429	2	O8RFD2	O8RFD2 fusobacteri
40	557	20.0	434	2	O820D1	O820D1 nitrosomona
41	552.5	19.8	429	2	O7P659	O7P659 fusobacteri
42	546	19.6	378	2	O9BS99	O9BS99 homo sapien
43	542.5	19.5	426	1	APER_MYCLE	O50022 mycobacteri
44	538.5	19.3	429	2	O6SG68	O6SG68 uncultured
45	502.5	18.0	428	2	O6AAB7	O6AAB7 propionibac

#### ALIGNMENTS

##### RESULT 1

ID Q9H85 PRELIMINARY, PRT, 535 AA.  
AC Q9H85;  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)  
DE Related to antipeptidase ysc1, vacuolar.  
GN Name=B13020.180;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Altm V., Hohlmeier J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.,  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AL451015, CAC18208.1, -;  
DR GO: GO:0005773, C:vacuole, IEA.  
DR GO: GO:0004250, F:aminopeptidase I activity, IEA.  
DR GO: GO:0006508, P:proteolysis and peptidolysis, IEA.  
DR Interpro: IPR001948; Peptidase M18.  
DR Pfam: PF02127; Peptidase M18; I.  
DR PRINTS: PR00932; AMINOPEPTASE.  
KW Aminopeptidase.  
SQ SEQUENCE 535 AA; 58137 MW; 469F22A535396798 CRC64;  
Alignment Scores:  
Pred. No.: 4.3e-102  
Score: 1528.00  
Percent Similarity: 72.35%  
Best Local Similarity: 57.77%  
Query Match: 54.89%  
Matches: 305  
Conservative: 77  
Mismatch: 119  
Indels: 27  
Gaps: 5



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 11, 2005, 05:12:35 ; Search time 125 Seconds  
(without alignments)  
9467.694 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602  
Perfect score: 2784  
Sequence: 1 atgaccaaagagagtcctc.....caagaattcgacagatttc 1530

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p\_model -DEV=xl  
-O=/cgn\_1/USPTO\_spool\_p/US1066958/runat\_10092005\_173354\_1967/app\_query.fasta\_1.1671  
-DB=4 Geneseq\_16Dec04 -QPMT=fastan -SUPFIX=top20.rag -MINMATCH=0.1 -LOOPT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1066958 @CGN\_1\_1\_154 @runat\_10092005\_173354\_1967 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEDIRECTORY -NEG SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: Geneseq\_16Dec04:\*  
2: geneseqp1980s:\*  
3: geneseqp1990s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2003bs:\*  
9: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2626	94.3	510	6	AAO19562 A nidulans
2	2112.5	75.9	519	6	AAO19561 A nidulans
3	1013.5	36.4	514	2	AAB05535 Yeast ami
4	821.5	29.5	487	3	AAB26490 Soybean a
5	746	26.8	485	4	AAB93510 Human pro
6	739	26.5	485	4	AAB94381 Human pro
7	737.5	26.5	471	3	AAV29464 Human ami
8	737.5	26.5	471	3	AAO17722 Human pro
9	737.5	26.5	471	3	AAO17715 Human pro
10	737.5	26.5	475	7	ADB31713 Human 616

11	737.5	26.5	475	8	ABM80387
12	732.5	26.3	471	2	AAV29466 Mouse ami
13	688	24.7	449	8	AD116362 Human pro
14	681.5	24.5	485	7	ABO72890 Pseudomon
15	678	24.4	428	4	AAU05191 Pseudomon
16	640	23.0	470	3	AAO17723 C elegans
17	613.5	22.0	441	4	AAB93327 Human pro
18	560	20.1	407	8	ADP04194 Human col
19	559.5	20.1	459	2	AAV29465 Human ami
20	558	20.0	499	8	ADP04193 Human col
21	542.5	19.5	447	7	ADB74497 Mycobacte
22	502.5	18.0	428	4	AAU58862 Pseudonib
23	502.5	18.0	428	6	ABM55381 Pseudonib
24	499.5	17.9	315	3	AAAG23766 Arabidops
25	499.5	17.9	321	3	AAAG23765 Arabidops
26	498.5	17.9	420	4	AAAB79077 Corynebac
27	498.5	17.9	420	5	AAAB79402 Corynebac
28	498.5	17.9	441	4	AAAG91402 C glutam
29	356	12.8	267	4	AAAB93839 Human pro
30	174	6.2	163	3	AAAG23767 Arabidops
31	141	5.1	752	7	ABO72978 Pseudomon
32	124.5	4.5	94	5	ABP00629 Human ORF
33	120.5	4.3	925	7	ADM05204 Human pro
34	120.5	4.3	1688	4	ABM59327 Drosophi1
35	119	4.3	1093	7	ABO73081 Pseudomon
36	118.5	4.3	1938	6	ABAB98398 Streptomy
37	118	4.2	519	5	ABG93208 S. cerevi
38	118	4.2	519	8	AUS44008 Bacteri
39	117	4.2	346	8	ADN46314 Thermococ
40	116.5	4.2	1403	4	ABB60540 Drosophi1
41	116	4.2	649	4	AAU52702 Pseudonib
42	116	4.2	649	6	ABM49221 Pseudonib
43	113.5	4.1	1770	8	ADM97110 Mouse gna
44	113.5	4.1	2354	8	ADQ97650 Mouse can
45	112.5	4.0	629	7	ABO68404 Pseudomon

## ALIGNMENTS

RESULT 1	AAO19562 standard; protein; sig AA.
AAO19562	
ID	AAO19562 standard; protein; sig AA.
AC	AAO19562;
XX	
DT	23-OCT-2003 (revised)
DT	31-JAN-2003 (first entry)
XX	
DE	A nidulans aminopeptidase #2.
XX	
KW	Aminopeptidase; protein digestion; seasoning; soy sauce; cooking sauce.
XX	
OS	Emeticella nidulans.
XX	
PN	WO200277223-A1.
XX	
PD	03-OCT-2002.
XX	
PF	15-MAR-2002; 2002WO-JP002476.
XX	
PR	19-MAR-2001; 2001JP-00078930.
PR	26-SEP-2001; 2001JP-00293348.
XX	
PA	(AJIN ) AJINOMOTO CO INC.
XX	
PI	Kojibuchi K, Ninomiya D, Kojima M, Ueda Y, Maruyama J, Kitamoto K;
XX	WPI, 2003-018923/01.
DR	N-PSDB; AAL50169.
XX	
PT	Novel Aspergillus nidulans-originated aminopeptidase and encoded gene,
PT	capable of digesting hardly digestible peptides, applicable e.g. in
PT	producing natural seasoning materials containing protein hydrolyzates

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 11, 2005, 14:11:55 ; Search time 41.5 Seconds  
(without alignments)  
7094.537 Million cell updates/sec

Title: US-10-664-958-4\_COPY\_73\_1602  
Perfect score: 2784  
Sequence: 1 atgaccacaaaggatgctcct.....acaaggattcgagatttc 1530

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p model -DRV=xld  
-O=/cgn2.1/USPTO\_spool\_p/US10664958/runat\_10092005\_173355\_1981/app\_query\_fasta\_1.1671  
-DB=PIR\_79 -OPMT=faatan -SUFFIX=top20.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10664958 @CCN 1.1 40 @runat 10092005\_173355\_1981 -NCPU=6 -ICPU=3  
-NO\_MAP -LANG=OTHER -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGCLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1013.5	36.4	514	2	A33879
2	843	30.3	490	2	S48955
3	814.5	29.3	467	2	T38806
4	718	25.8	526	2	T48467
5	677	24.3	429	2	A83240
6	650.5	23.4	392	2	T43206
7	640	23.0	433	2	F96974
8	640	23.0	470	2	T15946
9	542.5	19.5	433	2	H87185
10	486	17.5	443	2	A70536
11	485.5	17.4	432	2	T36482
12	446.5	16.0	423	2	B70178
13	333	12.0	452	2	B72387
14	304.5	10.9	458	2	E70145

15	276.5	9.9	465	2	F97034	aspartyl aminopept
16	124.5	4.5	2152	2	T45583	hypothetical prote
17	120	4.3	528	2	I47141	gastric mucin (clo
18	118	4.2	519	1	S67286	probable aldehyde
19	117	4.2	1222	2	G59100	hypothetical prote
20	117	4.2	2232	2	T34434	hypothetical prote
21	115.5	4.1	553	2	T52362	hypothetical prote
22	115.5	4.1	1161	2	S57180	probable membrane
23	115	4.1	967	2	H85334	T2012.10 protein -
24	114.5	4.1	1299	2	T47182	hypothetical prote
25	114	4.1	424	2	A10465	probable glycerol-
26	114	4.1	1066	2	G84746	hypothetical prote
27	113.5	4.1	943	2	F70873	acetic acid hydrat
28	111	4.0	1138	2	G71554	probable transmem
29	109.5	3.9	915	2	T12526	hypothetical prote
30	108.5	3.9	4273	2	C69679	polyketide synthas
31	108	3.9	623	2	S46427	botulinum neurotox
32	107.5	3.9	422	2	AG3323	threonine dehydrat
33	107.5	3.9	565	2	T26161	hypothetical prote
34	107.5	3.9	850	2	AE1954	hypothetical prote
35	107.5	3.9	2090	2	S26058	probable transform
36	107.5	3.9	3535	2	E83641	probable hemagglut
37	106	3.8	354	2	E71193	probable operon pr
38	106	3.8	619	2	C70669	probable acyl-CoA
39	106	3.8	1169	2	S8181	flocculation prote
40	105.5	3.8	662	2	E97016	transketolase (imp
41	105.5	3.8	887	2	D82049	phosphoenolpyruvat
42	105.5	3.8	898	1	DBPPT4	DNA-directed DNA p
43	105.5	3.8	1097	2	S17308	leukemia inhibitor
44	105	3.8	2440	2	S39162	transcription coac
45	104.5	3.8	422	2	AH1694	threonine dehydrat

## ALIGNMENTS

RESULT 1  
A33879  
aminopeptidase ysc1 (EC 3.4.11.-) precursor, vacuolar - yeast (Saccharomyces cerevisia  
N/Altemate names: protein YK103c; protein YK1455  
C/Species: Saccharomyces cerevisiae  
C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text\_change 09-Jul-2004  
C/Accession: A33879; B33879; S39101; S37930; S25400  
R/Chang, Y.H.; Smith, J.A.  
J. Biol. Chem. 264, 6979-6983, 1989  
A/Title: Molecular cloning and sequencing of genomic DNA encoding aminopeptidase I fro  
A/Reference number: A33879; NCBI:89214116; PMID:2651436  
A/Accession: A33879  
A/Molecule type: DNA  
A/Residues: 1-514 <CHA>  
A/Cross-references: UNIPROT:P14904; EMBL:M25548; NID:G295617; PIDN:AAA34738.1; PID:G29  
A/Accession: B33879  
A/Molecule type: protein  
A/Residues: 46-63 <CH2>  
R/Cheret, G.; Pallier, C.; Valens, M.; Daigman-Fornier, B.; Fukuhara, H.; Bolotin-Fuku  
Yeast 9, 1259-1265, 1993  
A/Title: The DNA sequence analysis of the HAP4-LAP4 region on chromosome XI of Sacchar  
A/Reference number: S39094; NCBI:94152173; PMID:8109175  
A/Accession: S39101  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-514 <CH2>  
A/Cross-references: EMBL:X71133; NID:G411205; PIDN:CAA50454.1; PID:G411213  
A/Experimental source: strain S288C  
R/Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daigman-Fornier, B.; Pallier, C.; Fu  
submitted to the Protein Sequence Database, March 1994  
A/Reference number: S37920  
A/Accession: S37930  
A/Molecule type: DNA  
A/Residues: 1-514 <CHR>  
A/Cross-references: EMBL:Z28103; NID:G486172; PIDN:CAA81943.1; PID:G486173; MIPS:YK110  
A/Experimental source: strain S288C  
R/Cueva, R.; Garcia-Alvarez, N.; Suarez-Rendueles, P.  
FEBS Lett. 259, 125-129, 1989